

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 13:43:56 ; Search time 38 Seconds
(without alignments)
10.471 Million cell updates/sec

Title: US-09-699-224A-8
Perfect score: 28
Sequence: 1 DEXGLF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	27	96.4	45	6	5185254-1	Patent No. 5185254	Sequence 2, Appli
2	27	96.4	265	4	US-08-413-805-2	Sequence 4196, Ap	Sequence 2, Appli
3	27	96.4	297	4	US-09-107-532A-4196	Sequence 8157, Ap	Sequence 82, Appl
4	27	96.4	314	4	US-09-919-497-82	Sequence 8157, Ap	Sequence 82, Appl
5	27	96.4	356	4	US-09-543-681A-8157	Sequence 8157, Ap	Sequence 82, Appl
6	27	96.4	587	3	US-09-147-923-2	Sequence 31342, A	Sequence 2, Appli
7	27	96.4	597	4	US-09-252-991A-31342	Sequence 31342, A	Sequence 2, Appli
8	26	92.9	80	3	US-09-247-155-93	Sequence 93, Appl	Sequence 93, Appl
9	26	92.9	104	4	US-09-270-767-47871	Sequence 47871, A	Sequence 47871, A
10	26	92.9	130	4	US-09-107-532A-6103	Sequence 6103, Ap	Sequence 6103, Ap
11	26	92.9	260	4	US-09-489-039A-12019	Sequence 12019, A	Sequence 12019, A
12	26	92.9	313	4	US-09-710-279-1758	Sequence 1758, Ap	Sequence 1758, Ap
13	26	92.9	318	1	US-08-220-958-4	Sequence 4, Appli	Sequence 4, Appli
14	26	92.9	328	3	US-09-134-001C-3229	Sequence 3229, Ap	Sequence 3229, Ap
15	26	92.9	352	4	US-09-252-991A-19989	Sequence 19989, A	Sequence 19989, A
16	26	92.9	380	3	US-09-316-080-2	Sequence 2, Appli	Sequence 2, Appli
17	26	92.9	381	4	US-09-489-039A-8085	Sequence 8085, Ap	Sequence 8085, Ap
18	26	92.9	496	4	US-08-622-191-1	Sequence 1, Appli	Sequence 1, Appli
19	26	92.9	516	3	US-08-888-949-17	Sequence 17, Appl	Sequence 17, Appl
20	26	92.9	516	3	US-08-888-950-17	Sequence 17, Appl	Sequence 17, Appl
21	26	92.9	516	3	US-09-262-758-17	Sequence 17, Appl	Sequence 17, Appl
22	26	92.9	516	4	US-09-885-876-17	Sequence 17, Appl	Sequence 17, Appl
23	26	92.9	516	4	US-09-885-901-17	Sequence 17, Appl	Sequence 17, Appl
24	26	92.9	516	4	US-09-731-393-17	Sequence 9, Appli	Sequence 9, Appli
25	26	92.9	526	4	US-09-910-174B-9	Sequence 9, Appli	Sequence 9, Appli
26	26	92.9	526	4	US-09-620-461-9	Sequence 12, Appl	Sequence 12, Appl
27	26	92.9	529	1	US-08-484-815-12	Sequence 12, Appl	Sequence 12, Appl

28	92.9	529	3	US-08-888-949-12	Sequence 12, Appl
29	92.9	529	3	US-08-888-950-12	Sequence 12, Appl
30	92.9	529	3	US-09-262-758-12	Sequence 12, Appl
31	92.9	529	4	US-09-885-876-12	Sequence 12, Appl
32	92.9	529	4	US-09-351-150A-3	Sequence 3, Appli
33	92.9	529	4	US-09-885-901-12	Sequence 12, Appl
34	92.9	529	4	US-09-731-393-12	Sequence 12, Appl
35	92.9	529	5	PCT-US95-10284-12	Sequence 12, Appl
36	92.9	589	2	US-08-724-394A-1	Sequence 1, Appli
37	92.9	663	1	US-07-912-015-2	Sequence 2, Appli
38	92.9	762	4	US-09-579-365-2	Sequence 2, Appli
39	92.9	762	4	US-09-807-063-2	Sequence 2, Appli
40	92.9	947	4	US-09-719-190-2	Sequence 2, Appli
41	92.9	991	3	US-09-352-159-27	Sequence 27, Appl
42	92.9	991	3	US-09-352-168-27	Sequence 27, Appl
43	92.9	991	4	US-09-771-045B-27	Sequence 27, Appl
44	92.9	1196	3	US-09-352-159-31	Sequence 31, Appl
45	92.9	1196	3	US-09-352-168-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
5185254-1
; Patent No. 5185254
; APPLICANT: LINNENBACH, ALBAN
; TITLE OF INVENTION: GENE FAMILY OF TUMOR-ASSOCIATED ANTIGENS
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA:
; FILING DATE: 29-DEC-1988
; SEQ ID NO:1:
; LENGTH: 45
5185254-1

Query Match 96.4%; Score 27; DB 6; Length 45;
Best Local Similarity 83.3%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 20 DESGLF 25

RESULT 2
US-08-413-805-2
; Sequence 2, Application US/08413805
; Patent No. 6645498
; GENERAL INFORMATION:
; APPLICANT: LINNENBACH, ALBAN J.
; APPLICANT: KOPROWSKI, HILARY
; APPLICANT: HERLYN, DOROTHEE
; TITLE OF INVENTION: Soluble Variants of Type I Membrane
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: P.O. Box 457 Spring House Corporate Center
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,805
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 07/778,232
; FILING DATE: 18-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,216
; FILING DATE: 05-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST21BUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9200
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 265 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-413-805-2

Query Match          96.4%; Score 27; DB 4; Length 265;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DEXGLF 6
Db      100 DESGLF 105

RESULT 3
US-09-107-532A-4196
; Sequence 4196, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; TELEFAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 4196:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 297 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:

```

```

; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...297
; SEQUENCE DESCRIPTION: SEQ ID NO: 4196:
US-09-107-532A-4196

```

```

Query Match          96.4%; Score 27; DB 4; Length 297;
Best Local Similarity 83.3%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DEXGLF 6
Db      62 DEAGLF 67

```

```

RESULT 4
US-09-919-497-82
; Sequence 82, Application US/09919497
; Patent No. 6773883
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 82
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-82

```

```

Query Match          96.4%; Score 27; DB 4; Length 314;
Best Local Similarity 83.3%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DEXGLF 6
Db      100 DESGLF 105

```

```

RESULT 5
US-09-543-681A-8157
; Sequence 8157, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 8157
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-8157

```

```

Query Match          96.4%; Score 27; DB 4; Length 356;
Best Local Similarity 83.3%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DEXGLF 6
Db      132 DESGLF 137

```

```
RESULT 6
US-09-147-923-2
; Sequence 2, Application US/09147923
; Patent No. 6146863
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; APPLICANT: Pratt, Julie M.
; APPLICANT: Hodgson, John E.
; APPLICANT: Beattie, David T.
; APPLICANT: Lowe, Adrian M.
; APPLICANT: Lonetto, Michael A.
; APPLICANT: Nicholas, Richard O. L.
; APPLICANT: Dersiewicz, Robert L.
; TITLE OF INVENTION: hcd
; FILE REFERENCE: GM10104
; CURRENT APPLICATION NUMBER: US/09/147,923
; CURRENT FILING DATE: 1999-03-18
; EARLIER APPLICATION NUMBER: 60/060,983
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-147-923-2

Query Match          96.4%; Score 27; DB 3; Length 587;
Best Local Similarity 83.3%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
Db 334 DEAGLF 339

RESULT 7
US-09-252-991A-31342
; Sequence 31342, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31342
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31342

Query Match          96.4%; Score 27; DB 4; Length 597;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
Db 363 DEAGLF 368

RESULT 8
US-09-247-155-93
; Sequence 93, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
```

```
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET.021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 93
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -40...-1
US-09-247-155-93

Query Match          92.9%; Score 26; DB 3; Length 80;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
Db 73 DEDGLF 78
```

```
RESULT 9
US-09-270-767-47871
; Sequence 47871, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47871
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47871

Query Match          92.9%; Score 26; DB 4; Length 104;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
Db 97 DEGGLF 102
```

```
RESULT 10
US-09-107-532A-6103
; Sequence 6103, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
```

STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 6103:
SEQUENCE CHARACTERISTICS:
LENGTH: 130 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...130
SEQUENCE DESCRIPTION: SEQ ID NO: 6103:
US-09-107-532A-6103

Query Match 92.9%; Score 26; DB 4; Length 130;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|| |||
DB 33 DEIGLF 38

RESULT 11
US-09-489-039A-12019
Sequence 12019, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 12019
LENGTH: 260
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12019

Query Match 92.9%; Score 26; DB 4; Length 260;
Best Local Similarity 83.3%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|| |||
DB 8 DEIGLF 13

RESULT 12
US-09-710-279-1758
Sequence 1758, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUB480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1758
LENGTH: 313
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-09-710-279-1758

Query Match 92.9%; Score 26; DB 4; Length 313;
Best Local Similarity 83.3%; Pred. No. 88;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|| |||
DB 67 DELGLF 72

RESULT 13
US-08-220-958-4
Sequence 4, Application US/08220958
Patent No. 5459072
GENERAL INFORMATION:
APPLICANT: McKay, Larry
APPLICANT: Polzin, Kayla
TITLE OF INVENTION: FOOD-GRADE INTEGRATION VECTORS FOR
TITLE OF INVENTION: INDUSTRIAL BACTERIAL STRAINS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: MERCHANT & GOULD
STREET: 3100 No. 5459072west Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/220,958
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/841,642
FILING DATE: 25-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kowalchuk, Alan W.
REGISTRATION NUMBER: 31,535
REFERENCE/DOCKET NUMBER: 600.229-US-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-332-5300
TELEFAX: 612-332-9081

```
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Lactococcus lactis subsp lactis bv.
; ORGANISM: diacetylactis
; STRAIN: DRC3
; IMMEDIATE SOURCE:
; CLONE: Putative Nisin Resistance Gene
; US-08-220-958-4

Query Match          92.9%; Score 26; DB 1; Length 318;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
Db 50 DELGLF 55

RESULT 14
US-09-134-001C-3229
; Sequence 3229, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3229
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-3229

Query Match          92.9%; Score 26; DB 3; Length 328;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
Db 82 DELGLF 87

RESULT 15
US-09-252-991A-19989
; Sequence 19989, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19989
; LENGTH: 352
```

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-19989

Query Match          92.9%; Score 26; DB 4; Length 352;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
Db 230 DEDGLF 235

Search completed: January 3, 2005, 13:53:46
Job time : 39 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 13:52:27 ; Search time 142 Seconds
(without alignments)
15.200 Million cell updates/sec

Title: US-09-699-224A-8
Perfect score: 28
Sequence: 1 DEXGLF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1599051 seqs, 359727711 residues

Total number of hits satisfying chosen parameters: 1599051

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	96.4	13	14	US-10-325-694-60
2	27	96.4	13	14	Sequence 60, Appl
3	27	96.4	13	14	Sequence 61, Appl
4	27	96.4	13	14	US-10-325-694-61
5	27	96.4	13	14	US-10-325-694-62
6	27	96.4	123	14	US-10-325-694-63
7	27	96.4	132	16	US-10-156-761-10753
8	27	96.4	193	14	US-10-767-701-62789
9	27	96.4	244	15	US-10-106-698-6736
10	27	96.4	244	15	US-10-374-780A-1594
11	27	96.4	265	14	US-10-412-699B-1640
12	27	96.4	265	14	US-10-299-383-2
13	27	96.4	275	15	US-10-404-724-58
					Sequence 58, Appl
					Sequence 76844, A
					Sequence 1593, A

14	27	96.4	310	15	US-10-412-699B-1639	Sequence 1639, Ap
15	27	96.4	310	16	US-10-437-963-162911	Sequence 162911,
16	27	96.4	314	9	US-09-922-217-1110	Sequence 1110, Ap
17	27	96.4	314	9	US-09-919-497-82	Sequence 82, Appl
18	27	96.4	314	13	US-10-025-380-1110	Sequence 1110, Ap
19	27	96.4	314	14	US-10-205-823-403	Sequence 403, App
20	27	96.4	314	14	US-10-404-724-60	Sequence 60, Appl
21	27	96.4	314	14	US-10-341-434-73	Sequence 73, Appl
22	27	96.4	314	14	US-10-236-031B-12	Sequence 12, Appl
23	27	96.4	314	16	US-10-408-765A-38	Sequence 38, Appl
24	27	96.4	314	17	US-10-643-795A-115	Sequence 115, App
25	27	96.4	314	17	US-10-643-795A-133	Sequence 133, App
26	27	96.4	375	9	US-09-925-300-1596	Sequence 1596, Ap
27	27	96.4	375	14	US-10-106-698-4883	Sequence 4883, Ap
28	27	96.4	394	15	US-10-264-049-2328	Sequence 2328, Ap
29	27	96.4	680	9	US-09-738-626-4877	Sequence 4877, Ap
30	27	96.4	753	9	US-09-815-242-12462	Sequence 12462, A
31	27	96.4	753	15	US-10-282-122A-44186	Sequence 44186, A
32	27	96.4	771	9	US-09-815-242-5609	Sequence 5609, Ap
33	26	92.9	52	17	US-10-425-115-324179	Sequence 324179,
34	26	92.9	71	17	US-10-425-115-256131	Sequence 256131,
35	26	92.9	80	10	US-09-903-190-93	Sequence 93, Appl
36	26	92.9	80	11	US-09-978-360A-749	Sequence 749, App
37	26	92.9	83	15	US-10-424-599-157263	Sequence 157263,
38	26	92.9	90	17	US-10-425-115-311214	Sequence 311214,
39	26	92.9	101	17	US-10-425-115-234470	Sequence 234470,
40	26	92.9	106	16	US-10-437-963-150847	Sequence 150847,
41	26	92.9	107	17	US-10-856-499-868	Sequence 868, App
42	26	92.9	110	15	US-10-424-599-270021	Sequence 270021,
43	26	92.9	120	16	US-10-767-701-33969	Sequence 33969, A
44	26	92.9	143	15	US-10-335-977-8378	Sequence 8378, Ap
45	26	92.9	146	15	US-10-335-977-8379	Sequence 8379, Ap

ALIGNMENTS

RESULT 1

US-10-325-694-60
; Sequence 60, Application US/10325694
; Publication No. US20030148463A1
; GENERAL INFORMATION:
; APPLICANT: KUFER, PETER
; APPLICANTS: RAUM, TOBIAS
; TITLE OF INVENTION: NOVEL METHOD FOR THE PRODUCTION OF ANTI-HUMAN ANTIGEN RECEPTORS AND USES THEREOF
; FILE REFERENCE: 38164000
; CURRENT APPLICATION NUMBER: US/10/325,694
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/403,107
; PRIOR FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-10-325-694-60

Query Match 96.4%; Score 27; DB 14; Length 13;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEXGLF 6

Db 8 DESGLF 13

RESULT 2

US-10-325-694-61
; Sequence 61, Application US/10325694

```
; Publication No. US20030148463A1
; GENERAL INFORMATION:
; APPLICANT: KUFRER, PETER
; APPLICANT: RAUM, TOBIAS
; TITLE OF INVENTION: NOVEL METHOD FOR THE PRODUCTION OF ANTI-HUMAN ANTIGEN
; TITLE OF INVENTION: RECEPTORS AND USES THEREOF
; FILE REFERENCE: 38164000
; CURRENT APPLICATION NUMBER: US/10/325,694
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/403,107
; PRIOR FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-10-325-694-61

Query Match          96.4%; Score 27; DB 14; Length 13;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEXGLF 6
Db      6 DESGLF 11

RESULT 3
US-10-325-694-62
; Sequence 62, Application US/10325694
; Publication No. US20030148463A1
; GENERAL INFORMATION:
; APPLICANT: KUFRER, PETER
; APPLICANT: RAUM, TOBIAS
; TITLE OF INVENTION: NOVEL METHOD FOR THE PRODUCTION OF ANTI-HUMAN ANTIGEN
; TITLE OF INVENTION: RECEPTORS AND USES THEREOF
; FILE REFERENCE: 38164000
; CURRENT APPLICATION NUMBER: US/10/325,694
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/403,107
; PRIOR FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-10-325-694-62

Query Match          96.4%; Score 27; DB 14; Length 13;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEXGLF 6
Db      4 DESGLF 9

RESULT 4
US-10-325-694-63
; Sequence 63, Application US/10325694
; Publication No. US20030148463A1
; GENERAL INFORMATION:
; APPLICANT: KUFRER, PETER
; APPLICANT: RAUM, TOBIAS
; TITLE OF INVENTION: NOVEL METHOD FOR THE PRODUCTION OF ANTI-HUMAN ANTIGEN
; TITLE OF INVENTION: RECEPTORS AND USES THEREOF
; FILE REFERENCE: 38164000
```

```
; CURRENT APPLICATION NUMBER: US/10/325,694
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/403,107
; PRIOR FILING DATE: 1999-10-14
; NUMBER OF SEQ ID NOS: 152
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-10-325-694-63
```

```
Query Match          96.4%; Score 27; DB 14; Length 13;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEXGLF 6
Db      2 DESGLF 7
```

```
RESULT 5
US-10-156-761-10753
; Sequence 10753, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10753
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10753
```

```
Query Match          96.4%; Score 27; DB 14; Length 123;
Best Local Similarity 83.3%; Pred. No. 1-2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEXGLF 6
Db      39 DEAGLF 44
```

```
RESULT 6
US-10-767-701-62789
; Sequence 62789, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
```


SEQ ID NO 62789
LENGTH: 132
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: 18065567.pep
US-10-767-701-62789

Query Match 96.4%; Score 27; DB 16; Length 132;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|| || ||
DB 119 DEAGLF 124

RESULT 7
US-10-106-698-6736
Sequence 6736, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.

TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005PI
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 6736
LENGTH: 193

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC_FEATURE
LOCATION: (3)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC_FEATURE
LOCATION: (21)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC_FEATURE
LOCATION: (22)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: MISC_FEATURE
LOCATION: (189)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6736

Query Match 96.4%; Score 27; DB 14; Length 193;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|| || ||
DB 88 DESGLF 93

RESULT 8
US-10-374-780A-1594
Sequence 1594, Application US/10374780A
Publication No. US20040019927A1
GENERAL INFORMATION:
APPLICANT: Sherman, Bradley K
APPLICANT: Riemann, Jose Luis

APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creelman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddle, James
APPLICANT: Broun, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omaira
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: MBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: 09/837,944
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
NUMBER OF SEQ ID NOS: 2906
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1594
LENGTH: 244

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Orthologous to G1417

US-10-374-780A-1594

Query Match 96.4%; Score 27; DB 15; Length 244;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|| || ||
DB 26 DEAGLF 31

RESULT 9
US-10-412-699B-1640
Sequence 1640, Application US/10412699B
Publication No. US20040045049A1
GENERAL INFORMATION:
APPLICANT: Mendel Biotechnology, Inc.
APPLICANT: Zhang, James
APPLICANT: Fromm, Michael E.
APPLICANT: Heard, Jacqueline E.
APPLICANT: Riemann, Jose Luis
APPLICANT: Adam, Luc J.
APPLICANT: Broun, Pierre E.
APPLICANT: Pineda, Omaira
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddle, James S.
APPLICANT: Yu, Guo-Liang
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Samaha, Raymond R.
APPLICANT: Pilgrim, Marsha L.

```

; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1640
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-412-699B-1640

Query Match          96.4%; Score 27; DB 15; Length 244;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEXGLF 6
DB      26 DEAGLF 31

RESULT 10
US-10-299-383-2
; Sequence 2, Application US/10299383
; Publication No. US2003010393A1
; GENERAL INFORMATION:
; APPLICANT: Linnenbach, Alban J.
;           Koprowski, Hilary
;           Herlyn, Dorothee
; TITLE OF INVENTION: Soluble Variants of Type I Membrane
;                   Proteins, and Methods of Using Them
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: P.O. Box 457 Spring House Corporate Center
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/299,383
; FILING DATE: 19-NO. US2003010393A1-2002

```

```

; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,805
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/778,232
; FILING DATE: 18-OCT-1991
; APPLICATION NUMBER: US 08/148,216
; FILING DATE: 05-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST21BUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-5200
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 265 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-299-383-2

```

```

Query Match          96.4%; Score 27; DB 14; Length 265;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEXGLF 6
DB      100 DESGLF 105

```

```

RESULT 11
US-10-404-724-58
; Sequence 58, Application US/10404724
; Publication No. US20030203447A1
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Arnold H.
; TITLE OF INVENTION: Methods and Materials For Increasing Expression of Recombinant
;                   Polypeptides
; FILE REFERENCE: 13698US01
; CURRENT APPLICATION NUMBER: US/10/404,724
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/368,530
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-404-724-58

```

```

Query Match          96.4%; Score 27; DB 14; Length 265;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEXGLF 6
DB      100 DESGLF 105

```

```

RESULT 12
US-10-282-122A-76844
; Sequence 76844, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari

```

```
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76844
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Ureaplasma urealyticum
US-10-282-122A-76844

Query Match          96.4%; Score 27; DB 15; Length 275;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEXGLF 6
DB      231 DETGLF 236

RESULT 13
US-10-374-780A-1593
; Sequence 1593, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James E
; APPLICANT: Broun, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25

; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76844
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Ureaplasma urealyticum
US-10-282-122A-76844

Query Match          96.4%; Score 27; DB 15; Length 275;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DEXGLF 6
DB      231 DETGLF 236

RESULT 14
US-10-412-699B-1639
; Sequence 1639, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
```

```
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1639
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-412-699B-1639
```

```
Query Match          96.4%; Score 27; DB 15; Length 310;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 DEXGLF 6
        |||||
Db      26 DEAGLF 31
```

```
RESULT 15
US-10-437-963-162911
; Sequence 162911, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 162911
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61957C.1.pap
US-10-437-963-162911
```

```
Query Match          96.4%; Score 27; DB 16; Length 310;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 DEXGLF 6
        |||||
Db      26 DEAGLF 31
```

Search completed: January 3, 2005, 14:04:37
Job time : 143 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 13:42:26 ; Search time 191 Seconds
(without alignments)
18.075 Million cell updates/sec

Title: US-09-699-224A-8
Perfect score: 28
Sequence: 1 DEXGLF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	27	96.4	123	2	Q82IE2	Q82ie2 streptomyc
2	27	96.4	125	2	Q9X953	Q9x953 streptomyc
3	27	96.4	135	2	Q6HY09	Q6hy09 bacillus an
4	27	96.4	146	2	O27265	O27265 methanobact
5	27	96.4	159	2	Q9FAS4	Q9fas4 vibrio para
6	27	96.4	231	2	Q8G1K1	Q8g1k1 brucella su
7	27	96.4	231	2	Q8YGC1	Q8ygc1 brucella me
8	27	96.4	259	1	UPK3_BACC1	P62463 bacillus ce
9	27	96.4	259	1	UPK3_BACCR	Q810p1 bacillus ce
10	27	96.4	259	2	Q6HI21	Q6hi21 bacillus th
11	27	96.4	259	2	AAS41644	Aas41644 bacillus
12	27	96.4	275	1	PPG_UREPA	Q9pg76 ureaplasma
13	27	96.4	281	1	Y818_PRAE	Q8zye1 pyrobaculum
14	27	96.4	283	2	Q9ZU57	Q9zu57 arabidopsis
15	27	96.4	310	2	Q94D31	Q94d31 oryza sativ
16	27	96.4	314	1	TTD1_HUMAN	P16422 h tumor-asp
17	27	96.4	314	2	Q6FG26	Q6fg26 homo sapien
18	27	96.4	314	2	Q6RG49	Q6rg49 homo sapien
19	27	96.4	322	2	Q848P5	Q848p5 uncultured
20	27	96.4	324	2	Q9RL47	Q9rl47 streptomyc
21	27	96.4	342	2	Q8CUY1	Q8cuy1 oceanobacil
22	27	96.4	357	1	RF1_CHLTE	Q8kg45 chlorobium
23	27	96.4	359	2	Q6SHD1	Q6shd1 uncultured
24	27	96.4	359	2	AAR37689	Aar37689 unculture
25	27	96.4	375	2	Q988B4	Q988b4 rhizobium 1
26	27	96.4	388	2	Q72SH0	Q72sh0 leptospira
27	27	96.4	388	2	Q8F348	Q8f348 leptospira
28	27	96.4	388	2	AAS70012	Aas70012 leptospir
29	27	96.4	399	2	O6VYP3	O6vyp3 oryza sativ
30	27	96.4	399	2	BAC99825	Bac99825 oryza sat
31	27	96.4	415	2	Q9AD10	Q9ad10 streptomyc

32	27	96.4	430	2	Q7SEY0	Q7sey0 ashbya goss
33	27	96.4	430	2	AAS50314	Aas50314 ashbya go
34	27	96.4	442	1	JRKL_HUMAN	Q9y4a0 homo sapien
35	27	96.4	465	2	Q98KK4	Q98kk4 rhizobium 1
36	27	96.4	469	2	Q8U6R8	Q8ue68 agrobacteri
37	27	96.4	492	2	Q8C0U8	Q8cou8 mus musculu
38	27	96.4	498	2	Q6Q8A2	Q6q8a2 nicotiana t
39	27	96.4	498	2	Q6Q8A3	Q6q8a3 nicotiana t
40	27	96.4	498	2	AAS60195	Aas60195 nicotiana
41	27	96.4	498	2	AAS60196	Aas60196 nicotiana
42	27	96.4	520	1	JRKY_HUMAN	Q7564 homo sapien
43	27	96.4	556	2	Q6N011	Q6n011 homo sapien
44	27	96.4	556	2	Q8ZS90	Q8zsa0 homo sapien
45	27	96.4	556	2	CAE45862	Caes45862 homo sapi

ALIGNMENTS

RESULT 1
Q82IE2
ID Q82IE2 PRELIMINARY; PRT; 123 AA.
AC Q82IE2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative WhiB-family transcriptional regulator; putative role in cell
DE cycle control.
GN Name=whi1; OrderedLocusNames=SAV3216;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005034; BAC70927.1; -
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003482; WhiB.
DR Pfam; PF02467; WhiB; 1.
KW Complete proteome.
SQ SEQUENCE 123 AA; 13659 MW; 17C2C3587C6C0D26 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 123;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEXGLF 6
|||
Db 39 DEAGLF 44

RESULT 2
Q8X953 PRELIMINARY; PRT; 125 AA.
ID Q9X953

Q9X953;
 AC 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein whlI (Hypothetical protein SCO5046).
 GN Name=whlI; OrderedLocusNames=SCO5046; ORFNames=SK7.19;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A;
 RX MEDLINE=20170666; PubMed=10708372;
 RA Soliveri J.A., Gomez J., Bishai W.R., Chater K.F.;
 RT "Multiple paralogous genes related to the Streptomyces coelicolor
 RT developmental regulatory gene whlB are present in Streptomyces and
 RT other actinomycetes.";
 RL Microbiology 146:333-343 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A;
 RA Soliveri J.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.-W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147 (2002).
 DR EMBL; AJ239088; CAB43033.1; -;
 DR EMBL; AL939122; CAC05891.1; -;
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR003482; Whib.
 DR Pfam; PF02467; Whib; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 125 AA; 13917 MW; A3B341D3E12D12EA CRC64;
 Query Match 96.4%; Score 27; DB 2; Length 125;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DEXGLF 6
 Db 39 DEAGLF 44

RESULT 3
 Q6HY09
 ID G6HY09 PRELIMINARY; PRT; 135 AA.
 AC G6HY09;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE Bacitracin resistance protein, N-terminal.
 GN OrderedLocusNames=BAS2519;
 OS Bacillus anthracis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1392;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sterne;

RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
 RA Richardson P., Rubin E., Tice H.,
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AS017225; AAT54830.1; -;
 DR InterPro; IPR003824; BACA.
 DR Pfam; PF02673; BACA; 1.
 SQ SEQUENCE 135 AA; 15525 MW; C52A5B51D68A48C2 CRC64;
 Query Match 96.4%; Score 27; DB 2; Length 135;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DEXGLF 6
 Db 37 DEAGLF 42

RESULT 4
 O27265
 ID O27265 PRELIMINARY; PRT; 146 AA.
 AC O27265;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Conserved protein.
 GN OrderedLocusNames=MTH1197;
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT deltaH: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155 (1997).
 DR EMBL; AE000888; AAB85686.1; -;
 DR PIR; F69026; F69026.
 KW Complete proteome.
 SQ SEQUENCE 146 AA; 16248 MW; 82FD4036C13C696C CRC64;
 Query Match 96.4%; Score 27; DB 2; Length 146;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DEXGLF 6
 Db 117 DETGLF 122

RESULT 5
 Q9FAS4
 ID Q9FAS4 PRELIMINARY; PRT; 159 AA.
 AC Q9FAS4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE UreE.
 GN Name=ureE;
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=670;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=TH3996;
RX MEDLINE=20448934; PubMed=10992480;
RA Park K.S., Iida T., Yamaichi Y., Oyagi T., Yamamoto K., Honda T.;
RT "Genetic characterization of DNA region containing the trh and ure
RL genes of Vibrio parahaemolyticus,"
RN Infect. Immun. 68:5742-5748(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TH3996;
RA Park K., Iida T.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038238; BAB13789.1; -.
DR HSSP; P18317; IGMW.
DR GO; GO:0016151; P:nickel ion binding; IEA.
DR GO; GO:0006461; P:protein complex assembly; IEA.
DR GO; GO:0019627; P:purea metabolism; IEA.
DR InterPro; IPR007864; UreC.
DR InterPro; IPR004029; UreN.
DR Pfam; PF05194; UreE_C; 1.
DR Pfam; PF02814; UreE_N; 1.
SQ SEQUENCE 159 AA; 17592 MW; 362146808B5440F7 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 159;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 45 DEAGLF 50

RESULT 6
Q8G1K1 PRELIMINARY; PRT; 231 AA.
AC Q8G1K1
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BR0714;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22847741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.B., Lindler L.E., Hailing S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014377; AAN29643.1; -.
DR TIGR; BR0714; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 231 AA; 25232 MW; B210D1BBA1752831 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 231;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 102 DETGLF 107

RP SEQUENCE FROM N.A.
RC STRAIN=TH3996;
RA Park K., Iida T.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038238; BAB13789.1; -.
DR HSSP; P18317; IGMW.
DR GO; GO:0016151; P:nickel ion binding; IEA.
DR GO; GO:0006461; P:protein complex assembly; IEA.
DR GO; GO:0019627; P:purea metabolism; IEA.
DR InterPro; IPR007864; UreC.
DR InterPro; IPR004029; UreN.
DR Pfam; PF05194; UreE_C; 1.
DR Pfam; PF02814; UreE_N; 1.
SQ SEQUENCE 159 AA; 17592 MW; 362146808B5440F7 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 159;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 45 DEAGLF 50

RESULT 6
Q8G1K1 PRELIMINARY; PRT; 231 AA.
AC Q8G1K1
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BR0714;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22847741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.B., Lindler L.E., Hailing S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014377; AAN29643.1; -.
DR TIGR; BR0714; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 231 AA; 25232 MW; B210D1BBA1752831 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 231;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 102 DETGLF 107

```

```

RESULT 7
Q8YGC1 PRELIMINARY; PRT; 231 AA.
ID Q8YGC1
AC Q8YGC1
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE INSERTION SEQUENCE IS21 PUTATIVE ATP-BINDING PROTEIN.
GN OrderedLocusNames=BMEI1238;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Hasselkorn R., Kyripides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis,"
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009563; AAL52419.1; -.
DR PIR; AH3406; AH3406.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 231 AA; 25232 MW; B210COAFADF135F1 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 231;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 102 DETGLF 107

RESULT 8
UPK3_BACC1 STANDARD; PRT; 259 AA.
ID UPK3_BACC1
AC P62463;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Putative undecaprenol kinase 3 (EC 2.7.1.66) (Bacitracin resistance
DE protein 3).
GN Name=upk3; Synonyms=bacA3; OrderedLocusNames=BCE2732;
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Rasko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RT adaptations and a large plasmid related to Bacillus anthracis pX01,"
RL Nucleic Acids Res. 32:977-988(2004).
CC -!- FUNCTION: Probably phosphorylates undecaprenol to undecaprenyl
CC phosphate. Confers resistance to bacitracin (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + undecaprenol = ADP + undecaprenyl
CC phosphate.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- MISCELLANEOUS: Bacitracin is thought to be involved in inhibition
CC of peptidoglycan synthesis by sequestering undecaprenyl
CC diphosphate reducing the pool of lipid carrier available.
CC -!- SIMILARITY: Belongs to the upk family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE017272; AAS41644.1; -;
DR TIGR; BCE2732; -;
DR HAMAP; MF_01006; -; 1;
KW Antibiotic resistance; Complete proteome; Kinase; Transferase;
KW Transmembrane.
FT TRANSMEM 7 29 Potential.
FT TRANSMEM 39 58 Potential.
FT TRANSMEM 70 92 Potential.
FT TRANSMEM 102 119 Potential.
FT TRANSMEM 140 162 Potential.
FT TRANSMEM 172 194 Potential.
FT TRANSMEM 206 228 Potential.
FT TRANSMEM 238 255 Potential.
SQ SEQUENCE 259 AA; 29209 MW; 545969E859B8A16 CRC64;

Query Match 96.4%; Score 27; DB 1; Length 259;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEXGLF 6
Db 37 DEAGLF 42

RESULT 9

UPK3_BACCR STANDARD; PRT; 259 AA.
AC Q81CPL;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Putative undecaprenol kinase 3 (EC 2.7.1.66) (Bacitracin resistance protein 3).
GN Nameupk3; Synonym=bacA3; OrderedLocusNames=BC2711;
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA Kapural V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Pusch G., Haseelkorn R., Fonstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.C.;
RT Genome sequence of Bacillus cereus and comparative analysis with
RT Bacillus anthracis.";
RL Nature 423:97-91(2003).

CC -1- FUNCTION: Probably phosphorylates undecaprenol to undecaprenyl
CC phosphate. Confers resistance to bacitracin (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + undecaprenol = ADP + undecaprenyl
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- MISCELLANEOUS: Bacitracin is thought to be involved in inhibition
CC of peptidoglycan synthesis by sequestering undecaprenyl
CC diphosphate reducing the pool of lipid carrier available.
CC -1- SIMILARITY: Belongs to the upk family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE017006; AAP09665.1; -;

DR HAMAP; MF_01006; -; 1;
DR InterPro; IPR003824; BACA.
DR Pfam; PF02673; BACA; 1.
KW Antibiotic resistance; Complete proteome; Kinase; Transferase;
KW Transmembrane.
FT TRANSMEM 7 29 Potential.
FT TRANSMEM 39 58 Potential.
FT TRANSMEM 70 92 Potential.
FT TRANSMEM 102 120 Potential.
FT TRANSMEM 172 194 Potential.
FT TRANSMEM 209 231 Potential.
FT TRANSMEM 238 255 Potential.
SQ SEQUENCE 259 AA; 29186 MW; 24435D0AFF4C8D3C CRC64;

Query Match 96.4%; Score 27; DB 1; Length 259;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEXGLF 6
Db 37 DEAGLF 42

RESULT 10

Q6HI21 PRELIMINARY; PRT; 259 AA.
AC Q6HI21;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Probable undecaprenol kinase (Bacitracin resistance protein) (EC 2.7.1.66).
GN Name=bacA; ORFNames=BT9727.2479;
OS Bacillus thuringiensis serovar konkukian str. 97-27.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=281309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97-27;
RA Bretton T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Tice H.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017355; AAT60038.1; -;
DR InterPro; IPR003824; BACA.
DR Pfam; PF02673; BACA; 1.
KW Kinase; Transferase.
SQ SEQUENCE 259 AA; 29058 MW; 36101D761CD9EA8E CRC64;

Query Match 96.4%; Score 27; DB 2; Length 259;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEXGLF 6
Db 37 DEAGLF 42

RESULT 11

AAS41644 PRELIMINARY; PRT; 259 AA.
ID AAS41644;
AC AAS41644;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 04-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE Undecaprenol kinase, putative (EC 2.7.1.66).
GN BCE2732.
OS Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=22523;
RN [1]
RP SEQUENCE FROM N.A.

RX PubMed=14960714;
RA Rasko D.A., Ravel J., Oskstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic
RT adaptations and a large plasmid related to *Bacillus anthracis* pXO1.";
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AE017272; AAS41644.1; -.
DR TIGR; BCE2732; -.
KW Kinase; Transferase.
SQ SEQUENCE 259 AA; 29209 MW; 545969E59BB8A16 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 259;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 37 DEAGLF 42

RESULT 12
FPG_UREPA
ID_PPG_UREPA STANDARD; PRT; 275 AA.
AC Q9PQ76;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23) (Fapy-DNA
DE glycosylase).
GN Name=mutM; Synonyms=fpg; OrderedLocusNames=UU413;
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serovar 3;
RX MEDLINE=20500219; PubMed=11048724; DOI=10.1038/35037619;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen *Ureaplasma*
RT *urealyticum*.";
RL Nature 407:757-762(2000).
CC -!- FUNCTION: This enzyme may play a significant role in processes
CC leading to recovery from mutagenesis and/or cell death by
CC alkylating agents (By similarity).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of DNA containing ring-opened N(7)-
CC methylguanine residues, releasing 2,6-diamino-4-hydroxy-5-(N-
CC methyl)formamidopyrimidine.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SIMILARITY: Belongs to the FPG family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE002138; AAF30824.1; -.
DR HSSP; P42371; INNJ.
DR HAMAP; MF_00103; -.
DR InterPro; IPR000214; Fapy_DNAG_Zn_BS.
DR InterPro; IPR000191; Fapy_DNA_glyco.
DR InterPro; IPR010979; Ribosomal_H2TH.
DR InterPro; IPR010663; ZF-FPG_ILKRS.
DR Pfam; PF01149; Fapy_DNA_glyco; 1.
DR Pfam; PF06831; H2TH; 1.
DR Pfam; PF06827; zf-FPG_ILKRS; 1.
DR ProDom; PD003680; Fapy_DNA_glyco; 1.
DR TIGRFAMs; TIGR00577; fpg; 1.
DR PROSITE; PS01242; FPG; 1.

KW Complete proteome; DNA repair; Glycosidase; Hydrolase; Zinc;
KW Zinc-finger.
FT ZN_FING 250 273 Potential.
SQ SEQUENCE 275 AA; 31993 MW; 9F4624E870ESAFA6A CRC64;

Query Match 96.4%; Score 27; DB 1; Length 275;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 231 DETGLF 236

RESULT 13
Y818_PYRAE
ID_Y818_PYRAE STANDARD; PRT; 281 AA.
AC Q8ZVE1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical UPF0103 protein PAE0818.
GN OrderedLocusNames=PAE0818;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum*
RT *aerophilum*.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC -!- SIMILARITY: Belongs to the UPF0103 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009787; AAL63052.1; -.
DR HAMAP; MF_00055; -.
DR InterPro; IPR002737; DUF52.
DR InterPro; IPR011057; Mss4_like.
DR Pfam; PF01875; UPF0103; 1.
DR ProDom; PD006364; DUF52; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 281 AA; 31033 MW; 85795161BF3E0EAD CRC64;

Query Match 96.4%; Score 27; DB 1; Length 281;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 217 DEAGLF 222

RESULT 14
Q3ZU57
ID_Q3ZU57 PRELIMINARY; PRT; 283 AA.
AC Q9ZU57;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative CENP-B/ARS binding protein-like protein.
GN Name=Atc2g06660;

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
 RA Barnstead M.B., Mason T.M., Bowman C.L., Renning C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.;
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RA Town C.D., Kaul S.;
 RN Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AC006161; AAD14510.1; -.
 DR PIR; B84479; B84479.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004875; CNBP-B.
 DR InterPro; IPR009057; Homeodomain_like.
 DR Pfam; PF03184; DDE; 1.
 SQ SEQUENCE 283 AA; 33068 MW; 8401D57F17BA63C4 CRC64;

Query Match 96.4%; Score 27; DB 2; Length 283;
 Best Local Similarity 83.3%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DEXGLF 6
 DB 81 DETGLF 86

RESULT 15

Q94D31 ID Q94D31 PRELIMINARY; PRT; 310 AA.
 AC Q94D31;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE P0712E02.25 protein (WRKY transcription factor 27) (P0700A11.6 protein).
 GN Names=P0712E02.25; Synonyms=P0700A11.6, WRKY27;
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Niiimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto M., Ando I., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
 RA Ikono M., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
 RA Waki K., Yamagata H., Yanane H., Yoshiki S., Yoshihara R., Yukawa K.,
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 RA Yano M., Jiang J., Gojobori T.;
 RT "The genome sequence and structure of rice chromosome 1.";
 RL Nature 420:312-316(2002).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX PubMed=15047897;
 RA Zhang Z.L., Xie Z., Zou X., Casaretto J., Ho T.H., Shen Q.J.;
 RT "A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells.";

RL Plant Physiol. 134:1500-1513(2004).
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Zhang Z.-L., Xie Z., Zou X., Casaretto J., Ho T.-h.D., Shen Q.J.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003492; BAB61861.1; -.
 DR EMBL; BK005030; DAA05092.1; -.
 DR EMBL; AP003300; BAB89907.1; -.
 DR Gramene; Q94D31; -.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR InterPro; IPR003657; WRKY.
 DR Pfam; PF03106; WRKY; 1.
 DR PROSITE; PS50811; WRKY; 1.
 SQ SEQUENCE 310 AA; 32557 MW; C06F9776CFB64161 CRC64;
 Query Match 96.4%; Score 27; DB 2; Length 310;
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DEXGLF 6
 DB 26 DEAGLF 31

Search completed: January 3, 2005, 13:52:20
 Job time : 193 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 13:43:11 ; Search time 38 Seconds
(without alignments)
15.192 Million cell updates/sec

Title: US-09-699-224A-8

Perfect score: 28
Sequence: 1 DEXGLF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: Piri.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	27	96.4	146	2 F63026	conserved hypotet
2	27	96.4	231	2 AH3406	insertion sequence
3	27	96.4	275	2 B82895	formamidopyrimidin
4	27	96.4	283	2 B84479	hypothetical prote
5	27	96.4	314	2 B48149	epithelial glycopr
6	27	96.4	442	2 JC5594	jerky gene protein
7	27	96.4	469	2 AG3139	flavin-containing
8	27	96.4	520	2 JE0225	JH8 protein - huma
9	27	96.4	605	2 E98148	hypothetical prote
10	27	96.4	694	2 E97630	hypothetical prote
11	27	96.4	694	2 AH2853	conserved hypotet
12	27	96.4	753	2 C89786	hypothetical prote
13	26	92.9	108	1 PVFG	parvalbumin beta -
14	26	92.9	111	2 A69932	hypothetical prote
15	26	92.9	143	2 G71880	hypothetical prote
16	26	92.9	143	2 F64634	hypothetical prote
17	26	92.9	161	2 S04765	LAT52 protein proc
18	26	92.9	174	2 F83581	heme d1 biosynthes
19	26	92.9	204	2 T47721	hypothetical prote
20	26	92.9	205	2 G97216	probable S-adenosy
21	26	92.9	212	2 F83917	hypothetical prote
22	26	92.9	235	2 AD3274	uroporphyrinogen-I
23	26	92.9	273	2 I39791	transcription regu
24	26	92.9	295	2 A64756	attaching and effa
25	26	92.9	313	2 G81145	hypothetical prote
26	26	92.9	314	1 A46489	pan-epithelial gly
27	26	92.9	318	2 A43746	nisin resistance p
28	26	92.9	322	2 G69946	phage-related prot
29	26	92.9	337	2 H64389	isocitrate dehydro

30	26	92.9	345	2 C84489	probable replicati
31	26	92.9	375	2 A87665	glycine cleavage s
32	26	92.9	381	2 B69722	queuine tRNA-ribos
33	26	92.9	383	1 S15624	E2 protein - human
34	26	92.9	385	2 T03003	enterohemolysin 1
35	26	92.9	390	2 T44324	hypothetical prote
36	26	92.9	417	2 JC5326	methicillin resist
37	26	92.9	496	2 S71205	hexokinase (EC 2.7
38	26	92.9	526	2 S70587	butyrophilin precu
39	26	92.9	526	2 A37821	butyrophilin - bov
40	26	92.9	646	2 S01351	type III site-spec
41	26	92.9	809	2 T18970	hypothetical prote
42	26	92.9	810	2 T19879	hypothetical prote
43	26	92.9	811	2 S17908	hypothetical prote
44	26	92.9	831	2 S62751	probable DNA-direc
45	26	92.9	1417	2 H90670	probable invasiv [

ALIGNMENTS

RESULT 1

F69026
conserved hypothetical protein MTH1197 - Methanobacterium thermoautotrophicum (strain Del
C)Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: F69026
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; F
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J, Bacteriol. 179, 7135-7155, 1997
A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: F69026
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-146 <MTH>
A:Cross-references: UNIPROT:O27265; GB:AE000888; GB:AE000666; NID:g2622304; PIDN:AA88568
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1197
A:Start codon: TTG

Query Match 96.4%; Score 27; DB 2; Length 146;
Best Local Similarity 83.3%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
DB 117 DETGLF 122

RESULT 2

AH3406
insertion sequence IS21 probable ATP-binding protein BMEI1238 [imported] - Brucella melit
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AH3406
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessc
Proc Natl Acad Sci U S A 99, 443-448, 2002
A>Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AB3252; PMID:11756688
A:Accession: AH3406
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <KUR>
A:Cross-references: UNIPROT:Q8YGC1; GB:AE008917; PIDN:AAL52419.1; PID:g17983222; GSPDB:G
C:Genetics:
A:Gene: BMEI1238
A:Map position: 1

Query Match 96.4%; Score 27; DB 2; Length 231;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
 |||||
 Db 102 DETGLF 107

RESULT 3
 B82895
 formamidopyrimidine-DNA glycosylase U413 [imported] - Ureaplasma urealyticum
 C/Species: Ureaplasma urealyticum
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C/Accession: B82895
 R/Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to GenBank, February 2000
 A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
 A/Reference number: A82870
 A/Accession: B82895
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-275 <GLA>
 A/Cross-references: GB:AE002138; GB:AF222894; PIDN:AAF30824.1; GSPDB:GN001
 A/Experimental source: serovar 3; biovar 1
 C/Genetics:
 A/Gene: fpg; U413
 A/Genetic code: SGC3
 C/Superfamily: formamidopyrimidine-DNA glycosidase

Query Match 96.4%; Score 27; DB 2; Length 275;
 Best Local Similarity 83.3%; Pred. No. 63;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
 |||||
 Db 231 DETGLF 236

RESULT 4
 B84479
 hypothetical protein At2g06660 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C/Accession: B84479
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: B84479
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-283 <STO>
 A/Cross-references: UNIPROT:Q9ZU57; GB:AE002093; NID:g4262216; PIDN:AAD14510.1; GSPDB:GN
 C/Genetics:
 A/Gene: At2g06660
 A/Map position: 2

Query Match 96.4%; Score 27; DB 2; Length 283;
 Best Local Similarity 83.3%; Pred. No. 65;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
 |||||
 Db 81 DETGLF 86

RESULT 5
 B48149
 epithelial glycoprotein antigen GA733-2 precursor - human
 N/Alternate names: adenocarcinoma-associated glycoprotein gp40; carcinoma-associated and

C/Species: Homo sapiens (man)
 C/Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C/Accession: B48149; A35907; A35995; A37377; A45988; A60298; A61569
 R/Linnenbach, A.J.; Seng, B.A.; Wu, S.; Robbins, S.; Scollon, M.; Pyrc, J.J.; Druck, T.;
 Mol. Cell. Biol. 13, 1507-1515, 1993
 A/Title: Retroposition in a family of carcinoma-associated antigen genes.
 A/Reference number: A48149; MUID:93180797; PMID:8382772
 A/Accession: B48149
 A/Molecule type: DNA
 A/Residues: 1-314 <LIN>
 A/Cross-references: UNIPROT:P16422; GB:M93036; NID:g182904; PIDN:AAB00775.1; PID:g182906
 R/Note: sequence extracted from NCBI backbone (NCBIN:125854, NCBIN:125857, NCBIN:125862,
 R/Zsala, S.; Froehlich, M.; Scollon, M.; Kasal, Y.; Stepiewski, Z.; Koprowski, H.; Liner
 Proc. Natl. Acad. Sci. U.S.A. 87, 3542-3546, 1990
 A/Title: Molecular cloning of cDNA for the carcinoma-associated antigen GA733-2.
 A/Reference number: A35907; MUID:90239051; PMID:2333300
 A/Accession: A35907
 A/Molecule type: mRNA
 A/Residues: 1-314 <SZ>
 A/Cross-references: GB:M33011; NID:g182895; PIDN:AAA35861.1; PID:g182896
 A/Experimental source: colorectal carcinoma cell line SW948
 R/Simon, B.; Podolsky, D.K.; Moldenhauer, G.; Isselbacher, K.J.; Gattoni-Celli, S.; Branc
 Proc. Natl. Acad. Sci. U.S.A. 87, 2755-2759, 1990
 A/Title: Epithelial glycoprotein is a member of a family of epithelial cell surface anti
 A/Reference number: A35995; MUID:90207276; PMID:2108441
 A/Accession: A35995
 A/Molecule type: mRNA
 A/Residues: 1-114, 'M', 116-314 <SIM>
 A/Cross-references: GB:M32306; NID:g181132; PIDN:AAA35723.1; PID:g181133
 A/Experimental source: colon carcinoma cell line HT-29
 R/Strnad, J.; Hamilton, A.E.; Beavers, L.S.; Gamboa, G.C.; Apelgren, L.D.; Taber, L.D.; f
 Cancer Res. 49, 314-317, 1989
 A/Title: Molecular cloning and characterization of a human adenocarcinoma/epithelial cell
 A/Reference number: A37377; MUID:89089570; PMID:2463074
 A/Accession: A37377
 A/Molecule type: mRNA
 A/Residues: 1-276, 'M', 278-314 <STR>
 A/Cross-references: GB:M32325; NID:g186775; PIDN:AAA36151.1; PID:g307091
 R/Perez, M.S.; Walker, L.E.
 J. Immunol. 142, 3662-3667, 1989
 A/Title: Isolation and characterization of a cDNA encoding the KSI/4 epithelial carcino
 A/Reference number: A45898; MUID:89235177; PMID:2469722
 A/Accession: A45898
 A/Molecule type: mRNA
 A/Residues: 1-276, 'M', 278-314 <PER>
 A/Cross-references: GB:M26481; NID:g619789; PIDN:AAA59543.1; PID:g619790
 R/Durbin, H.; Rodrigues, N.; Bodmer, W.F.
 Int. J. Cancer 45, 562-565, 1990
 A/Title: Further characterization, isolation and identification of the epithelial cell-s
 A/Reference number: A60298; MUID:90170210; PMID:2307544
 A/Accession: A60298
 A/Molecule type: protein
 A/Residues: 82-98 <DUR>
 A/Experimental source: normal colon mucosa
 R/Sportsman, J.R.; Taber, L.D.; Sliasz, M.C.; Apelgren, L.D.; Bumol, T.F.
 Biotechnol. Appl. Biochem. 10, 536-544, 1988
 A/Title: Isolation and characterization of the human adenocarcinoma- associated glycoprot
 A/Reference number: A61569; MUID:89166093; PMID:3069116
 A/Accession: A61569
 A/Molecule type: protein
 A/Residues: 82-97 <SPO>
 C/Superfamily: epithelial glycoprotein antigen GA733; EGF homology; thyroglobulin type I
 C/Keywords: glycoprotein; intestine; transmembrane protein
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-81/Domain: propeptide #status predicted <PRO>
 F:66-135/Domain: thyroglobulin type I repeat homology <THY1>
 F:82-314/Product: epithelial glycoprotein antigen GA733-2 #status predicted <MAT>
 F:82-265/Domain: extracellular #status predicted <EXT>
 F:266-288/Domain: transmembrane #status predicted <TM>
 F:289-314/Domain: intracellular #status predicted <INT>
 F:74,111,198/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.4%; Score 27; DB 2; Length 314;

```
Best Local Similarity 83.3%; Pred. No. 72;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
  |||||
Db 100 DESGLF 105

RESULT 6
JC5594
jerky gene protein homolog - human
C:Species: Homo sapiens (man)
C>Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 09-Jul-2004
C:Accession: JC5594
R:Zeng, Z.; Kyaw, H.; Gakenheimer, K.R.; Augustus, M.; Fan, P.; Zhang, X.; Su, K.; Carte
Biochem. Biophys. Res. Commun. 236, 389-395, 1997
A:Title: Cloning, mapping, and tissue distribution of a human homologue of the mouse jer
A:Reference number: JC5594; MUID:97382443; PMID:9240447
A:Accession: JC5594
A:Molecule type: mRNA
A:Residues: 1-442 <ZEN>
A:Cross-references: UNIPROT:Q9Y4A0; DDBJ:AF004715; NID:G2314828; PIDN:AAB65833.1; PID:G2
A:Note: it is uncertain whether Met-1 or Met-33 is the initiator
C:Comment: This protein functions as a nuclear regulatory protein.
C:Genetics:
A:Gene: hnmjg
A:Map position: 11q21

Query Match 96.4%; Score 27; DB 2; Length 442;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
  |||||
Db 95 DETGLF 100

RESULT 7
AG3139
flavin-containing monooxygenase [imported] - Agrobacterium tumefaciens (strain C58, Dupo
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AG3139
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG3139
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-469 <KUR>
A:Cross-references: UNIPROT:Q8U6R8; GB:AE008689; PIDN:AAL45533.1; PID:g17743245; GSPDB:G
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4739
A:Map position: linear chromosome

Query Match 96.4%; Score 27; DB 2; Length 469;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
  |||||
Db 310 DESGLF 315

RESULT 8
JE0225
JH8 protein - human
```

```
C:Species: Homo sapiens (man)
C>Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 17-Mar-1999
C:Accession: JE0225
R:Morita, R.; Miyazaki, E.; Fong, C.G.; Chen, X.N.; Korenberg, J.R.; Delgado-Esueta, A.V.
Biochem. Biophys. Res. Commun. 248, 307-314, 1998
A:Title: JH8, a gene highly homologous to the mouse jerky gene, maps to the region for c
A:Reference number: JE0225; MUID:98342085; PMID:9675132
A:Accession: JE0225
A:Molecule type: mRNA
A:Residues: 1-520 <MOR>
A:Experimental source: brain
A:Note: homologous to mouse jerky protein
C:Genetics:
A:Gene: JH8
A:Map position: 8q24.3

Query Match 96.4%; Score 27; DB 2; Length 520;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
  |||||
Db 139 DETGLF 144

RESULT 9
E98148
hypothetical protein AGR_L_284 [imported] - Agrobacterium tumefaciens (strain C58, Cereor
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: E98148
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: E98148
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-605 <KUR>
A:Cross-references: UNIPROT:Q8U6R8; GB:AE007870; PIDN:AAK88711.1; PID:g15158447; GSPDB:G
C:Genetics:
A:Gene: AGR_L_284
A:Map position: linear chromosome

Query Match 96.4%; Score 27; DB 2; Length 605;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
  |||||
Db 446 DESGLF 451

RESULT 10
E97630
hypothetical protein AGR_C_4105 [imported] - Agrobacterium tumefaciens (strain C58, Cereor
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: E97630
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: E97630
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-694 <KUR>
A:Cross-references: UNIPROT:Q8UD70; GB:AE007869; PIDN:AAK87998.1; PID:g15157412; GSPDB:G
C:Genetics:
A:Gene: AGR_C_4105
A:Map position: circular chromosome
```

```
Query Match      96.4%; Score 27; DB 2; Length 694;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels
```

Qy	1	DEXGLF	6
Db	576	DETGLF	581

RESULT 11

conserved hypothetical protein Atu2257 [imported] - *Agrobacterium tumefaciens* (strain C58) C:Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AH2853
R/Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.

A;Map position: circular chromosome

Qy	1	DEXGLF	6
Db	576	DETGLF	581

RESULT 12

C89766
hypothetical protein SA0224 [imported] - Staphylococcus aureus (strain N315)
C_Species: Staphylococcus aureus
C_Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C_Accession: C89766
C_C: Shiba, T.; Hattori, M.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
R.M.A., Kuroda, M.; Ohta, T.; Mizutani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.;
C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
C.Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A.Reference number: A89759; PMID:21311952; PMID:14118146

Query Match 96.4%; Score 27; DB 2; Length 753;
Best Local Similarity 83.3%;
Pred. NO. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels

Qy 1 DEXGLF 6
334 DEAGLF 339
db

RESULT 13

parvalbumin beta - edible frog
 C/Species: Rana esculenta (edible frog)
 C/Date: 30-Nov-1980 #sequence_revision 30-Nov-1980 #t
 C/Accession: A03052
 R/Capony, J.P.; Demalle, J.; Pina, C.; Pechere, J.F.

A;Accession: A03052
 A;Reference number: A03052; MUID:76022442; PMID:1080707
 A;Molecule type: protein
 A;Residues: 1-108 <CAP>
 A;Title: The amino-acid sequence of the most acidic major parvalbumin from frog muscle

Query Match 92.9%; Score 26; DB 1; Length 108;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	DEXGLF	6
Db	61	DELGLF	66

RESULT 14

A69932
 hypothetical protein yozM - Bacillus subtilis
 C-Species: Bacillus subtilis
 C-date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C-accession: A69932
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, C.; Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Nature 390, 249-256, 1997

Query Match 92.9%; Score 26; DB 2; Length 111;
Best Local Similarity 83.3%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 1; Indels

Qy	1	DEXGLF	6
Db	39	DEFGLE	44

RESULT 15

G71880
hypothetical protein jhp0852 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: G71880
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: G71880
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-143 <ARN>
A:Cross-references: UNIPROT:Q9ZKT3; GB:AE001514; GB:AE001439; NID:g4155418; PIDN:AAD0643
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0852

Query Match 92.9%; Score 26; DB 2; Length 143;
Best Local Similarity 83.3%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|||
DB 119 DERGLF 124

Search completed: January 3, 2005, 13:53:03
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2005, 13:39:06 ; Search time 156 Seconds
(without alignments)
13.797 Million cell updates/sec

Title: US-09-699-224A-8
Perfect score: 28
Sequence: 1 DEXGLF 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_238sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	96.4	13	AAW80859	Amino aci
2	27	96.4	13	AAW80858	Amino aci
3	27	96.4	13	AAW80856	Amino aci
4	27	96.4	13	AAW80857	Amino aci
5	27	96.4	66	ABP00014	Human ORF
6	27	96.4	76	AAO19722	Human GA7
7	27	96.4	106	AAI95248	Colorecta
8	27	96.4	107	ABG13551	Novel hum
9	27	96.4	126	AAO01704	Human pol
10	27	96.4	193	AAAG75962	Human col
11	27	96.4	240	ADP81219	Protein o
12	27	96.4	244	ADI43131	Plant tra
13	27	96.4	244	ADO03227	Thalacres
14	27	96.4	265	AAK36397	Truncated
15	27	96.4	265	ADL23185	Human sol
16	27	96.4	268	ADP81217	Protein o
17	27	96.4	275	ABU48920	Protein e
18	27	96.4	297	ADC94569	E. faeciu
19	27	96.4	310	ADI43130	Plant tra
20	27	96.4	310	ADO03226	Thalacres
21	27	96.4	313	AAO19719	Human GA7
22	27	96.4	313	ADL06553	Human tum
23	27	96.4	314	AAP90530	DNA encod
24	27	96.4	314	AAU57500	Propionib
25	27	96.4	314	AAU84292	Human end

26	27	96.4	314	6	ABM54019	Propionib
27	27	96.4	314	6	ABP55381	Human col
28	27	96.4	314	7	ADB70320	Tumour-as
29	27	96.4	314	7	ADB75579	Prostate
30	27	96.4	314	7	ADD45201	Human Pro
31	27	96.4	314	7	ADJ68232	Human hea
32	27	96.4	314	7	ADL23187	Human ful
33	27	96.4	314	8	ADU37099	Human mal
34	27	96.4	314	8	ADK70436	Respirato
35	27	96.4	314	8	ADK71156	Rhesus mo
36	27	96.4	314	8	ADK71158	Human Ep-
37	27	96.4	314	8	ADL06535	Human tum
38	27	96.4	314	8	ADN03783	Antipsoxi
39	27	96.4	314	8	ADP81221	Protein o
40	27	96.4	342	8	ADP81218	Bacteriol
41	27	96.4	356	7	ADF07872	Bacteriol
42	27	96.4	375	3	AAAB57018	Human pro
43	27	96.4	375	4	AAAG74109	Human col
44	27	96.4	383	4	ABB78746	Arabidops
45	27	96.4	387	4	ABG21997	Novel hum

ALIGNMENTS

RESULT 1
AAW80859
ID AAW80859 standard; peptide; 13 AA.
XX
AC AAW80859;
XX
DT 16-FEB-1999 (first entry)
XX
DE Amino acid sequence of the synthetic peptide 42.
XX
KW Synthetic peptide; human; receptor; antigen; tumour; auto-immune disease;
KW PCR; primer; graft rejection; allergy; inflammatory disease;
KW endocrine disease; degenerative disease.
XX
OS Synthetic.
XX
PN WO9846645-A2.
XX
PD 22-OCT-1998.
XX
PF 14-APR-1998; 98WO-EP002180.
XX
PR 14-APR-1997; 97EP-00106109.
XX
PA (KUFE/) KUFER P.
XX (RAUM/) RAUM T.
XX
PI Kufer P, Raum T;
XX
DR WPI; 1998-594564/50.
XX
PT Production of anti-human antigen receptors - by selecting a combination of functionally rearranged VH and VL immunoglobulin chains expressed from a recombinant vector.
XX
PS Example 5; Page 40; 84pp; English.
XX
CC This is the amino acid sequence of a synthetic peptide used in the method of the invention, involving the identification of receptors that can be used for targeting antigens in humans without being immunogenic themselves. Such receptors can be used for treating diseases such as tumours or auto-immune diseases, graft rejection after transplantation, infectious diseases by targeting cellular receptors as well as allergic, inflammatory, endocrine and degenerative diseases by targeting key molecules involved in the pathological process
XX
SQ Sequence 13 AA;

```

Query Match      96.4%; Score 27; DB 2; Length 13;
Best Local Similarity 83.3%; Pred. No. 9.4;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
   |||||
DB 2 DESGLF 7

RESULT 2
AAW80858
ID AAW80858 standard; peptide; 13 AA.
XX
AC AAW80858;
XX
DT 16-FEB-1999 (first entry)
XX
DE Amino acid sequence of the synthetic peptide 41.
XX
KW Synthetic peptide; human; receptor; antigen; tumour; auto-immune disease;
KW PCR; primer; graft rejection; allergy; inflammatory disease;
KW endocrine disease; degenerative disease.
XX
OS Synthetic.
XX
PN WO9846645-A2.
XX
PD 22-OCT-1998.
XX
PF 14-APR-1998; 98WO-EP002180.
XX
PR 14-APR-1997; 97EP-00106109.
XX
PA (KUFE/) KUFER P.
PA (RAUM/) RAUM T.
XX
PI Kufer P, Raum T;
XX
DR WPI; 1998-594564/50.
XX
PT Production of anti-human antigen receptors - by selecting a combination
PT of functionally rearranged VH and VL immunoglobulin chains expressed from
PT a recombinant vector.
XX
PS Example 5; Page 40; 84pp; English.
XX
CC This is the amino acid sequence of a synthetic peptide used in the method
CC of the invention, involving the identification of receptors that can be
CC used for targeting antigens in humans without being immunogenic
CC themselves. Such receptors can be used for treating diseases such as
CC tumours or auto-immune diseases, graft rejection after transplantation,
CC infectious diseases by targeting cellular receptors as well as allergic,
CC inflammatory, endocrine and degenerative diseases by targeting key
CC molecules involved in the pathological process
XX
SQ Sequence 13 AA;

Query Match      96.4%; Score 27; DB 2; Length 13;
Best Local Similarity 83.3%; Pred. No. 9.4;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
   |||||
DB 8 DESGLF 13

RESULT 4
AAW80857
ID AAW80857 standard; peptide; 13 AA.
XX
AC AAW80857;
XX
DT 16-FEB-1999 (first entry)
XX
DE Amino acid sequence of the synthetic peptide 40.
XX
KW Synthetic peptide; human; receptor; antigen; tumour; auto-immune disease;
KW PCR; primer; graft rejection; allergy; inflammatory disease;
KW endocrine disease; degenerative disease.
XX
OS Synthetic.
XX
PN WO9846645-A2.
XX
PD 22-OCT-1998.
XX
PF 14-APR-1998; 98WO-EP002180.
XX
PR 14-APR-1997; 97EP-00106109.

```

```
XX (KUFE/) KUFER P.
PA (RAUM/) RAUM T.
XX Kufer P, Raum T;
XX WPI; 1998-594564/50.
XX Production of anti-human antigen receptors - by selecting a combination
PT of functionally rearranged VH and VL immunoglobulin chains expressed from
PT a recombinant vector.
XX Example 5; Page 40; 84pp; English.
XX This is the amino acid sequence of a synthetic peptide used in the method
CC of the invention, involving the identification of receptors that can be
CC used for targeting antigens in humans without being immunogenic
CC themselves. Such receptors can be used for treating diseases such as
CC tumours or auto-immune diseases, graft rejection after transplantation,
CC infectious diseases by targeting cellular receptors as well as allergic,
CC inflammatory, endocrine and degenerative diseases by targeting key
CC molecules involved in the pathological process
XX Sequence 13 AA;
SQ
Query Match 96.4%; Score 27; DB 2; Length 13;
Best Local Similarity 83.3%; Pred. No. 9.4;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DEXGLF 6
DB 6 DESGLF 11
RESULT 5
ABP00014
ID ABP00014 standard; protein; 66 AA.
XX ABP00014;
XX 25-JUN-2002 (first entry)
XX Human ORFX protein sequence SEQ ID NO:10.
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX Homo sapiens.
XX WO200192523-A2.
XX 06-DEC-2001.
XX 29-MAY-2001; 2001WO-US010836.
XX 30-MAY-2000; 2000US-0206132P.
XX 29-AUG-2000; 2000US-0228716P.
XX (CURA-) CURAGEN CORP.
XX Shinketsu RA, Leach MD;
XX WPI; 2002-106308/14.
XX N-PSDB; ABN15766.
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
```

```
PT hyperproliferative disorders and autoimmune disorders.
XX Disclosure; SEQ ID NO 10; 1037pp; English.
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 66 AA;
SQ
Query Match 96.4%; Score 27; DB 5; Length 66;
Best Local Similarity 83.3%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DEXGLF 6
DB 45 DEAGLF 50
RESULT 6
AAO19722
ID AAO19722 standard; protein; 76 AA.
XX AAO19722;
XX 10-APR-2003 (first entry)
XX Human GA733-2 protein thyroglobulin domain.
XX Human; epithelial glycoprotein; mEGP; immune response suppression;
KW tumour-associated antigen; gene therapy; autoimmune disorder;
KW immunosuppressive; antidiabetic; neuroprotective; antirheumatic;
KW antiarthritic; dermatological; antiinflammatory; antianaemic.
XX Homo sapiens.
XX WO200288304-A2.
XX 07-NOV-2002.
XX 11-APR-2002; 2002WO-US011371.
XX 11-APR-2001; 2001US-0282946P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Eck SL;
XX WPI; 2003-167107/16.
XX N-PSDB; AB220997.
XX Modulating an undesirable immune response, e.g. transplant rejection, in
PT
```

PT patients, comprises administering a composition comprising a tumor-
 XX associated antigen molecule in a carrier to the patient.
 PS Claim 15; Fig 15; 87pp; English.
 XX
 CC The present invention relates to a method of modulating an undesirable
 CC immune response in a recipient host mammal to enhance tolerance in the
 CC recipient, which comprises administering at least one composition
 CC comprising a tumor-associated antigen molecule in a carrier to the
 CC recipient host mammal, in an amount to modulate the undesirable immune
 CC response. The method is useful in suppressing undesirable immune
 CC responses in humans which may lead to a disorder such as diabetes,
 CC lupus erythematosus, or pernicious anaemia. The nucleic acids and
 CC proteins identified in the invention may be used in screening assays to
 CC assess agents that modulate tumour-associated antigen mediated regulation
 CC of antigen recognition and immune suppression, as research tools to
 CC identify other proteins that are involved in antigen presentation by
 CC Major Histocompatibility Complex (MHC) Class II molecules, and in various
 CC diagnostic and therapeutic applications (e.g. in treating the above-
 CC mentioned diseases). The present sequence is the thyroglobulin domain of
 CC the human tumour-associated antigen GA733-2 sequence shown in the
 CC exemplification of the invention
 XX
 SQ Sequence 76 AA;
 Query Match 96.4%; Score 27; DB 6; Length 76;
 Best Local Similarity 83.3%; Pred. No. 62;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DEXGLP 6
 Db 38 DESGLF 43
 |||||
 |||||
 RESULT 7
 AAY95248
 ID AAY95248 standard; peptide; 106 AA.
 XX
 AC AAY95248;
 XX
 DT 29-AUG-2000 (first entry)
 XX
 DE Colorectal cancer antigen GA733-2.
 XX
 KW Colorectal cancer; antigen; GA733-2; virus infected plant; vaccine;
 KW pNF1/TVE; alfalfa mosaic virus; transcomplementation.
 XX
 OS Homo sapiens.
 XX
 PN WO200025574-A1.
 XX
 PD 11-MAY-2000.
 XX
 PF 29-OCT-1999; 99WO-US025566.
 XX
 PR 30-OCT-1998; 98US-0106221P.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Koprowski H, Yusibov V;
 XX
 DR WPI; 2000-365296/31.
 XX
 PT Producing full-length antibody in a host plant for use as vaccines,
 PT involves using recombinant viral vectors or transcomplementation systems.
 XX
 PS Example 1; Fig 1B; 87pp; English.
 XX
 CC This peptide is colorectal cancer antigen GA733-2, which is recognised by
 CC colorectal cancer associated antigen 17-1A. GA733-2 was fused with the
 CC coat protein of alfalfa mosaic virus (ALMV), and the fusion product was
 CC cloned into full-length RNA3 of ALMV to create pNF1/TVE. This construct

CC can be used to express GA733-2 in infected plants. In construct pNF2/TVE,
 CC non-full-length ALMV coat protein is used. The invention discloses novel
 CC methods for producing foreign polypeptides in a host plant using
 CC recombinant viral vectors, such as ALMV, and transcomplementation
 CC systems. The products expressed in the plant can be used as vaccine
 CC components or in other biomedical applications
 XX
 SQ Sequence 106 AA;
 Query Match 96.4%; Score 27; DB 3; Length 106;
 Best Local Similarity 83.3%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DEXGLP 6
 Db 85 DESGLF 90
 |||||
 |||||
 RESULT 8
 ABG13551
 ID ABG13551 standard; protein; 107 AA.
 XX
 AC ABG13551;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #13542.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS77738.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 43910; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 107 AA;

Query Match 96.4%; Score 27; DB 4; Length 107;
Best Local Similarity 83.3%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|||
Db 100 DESGLF 105

RESULT 9
AAO01704
ID AAO01704 standard; protein; 126 AA.

XX AAO01704;

DT 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 15596.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US004927.

PR 28-FEB-2000; 2000US-00515126.

PR 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

DR WPI: 2001-514838/56.

DR N-PSDB; AAI81635.

PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 15596; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 126 AA;

Query Match 96.4%; Score 27; DB 4; Length 126;
Best Local Similarity 83.3%; Pred. No. 11e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|||
Db 35 DESGLF 40

RESULT 10
AAG75962

ID AAG75962 standard; protein; 193 AA.

XX AAG75962;

DT 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:6726.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 4.

XX Homo sapiens.

XX WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US026524.

PR 29-SEP-1999; 99US-0157137P.

PR 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI: 2001-235357/24.

DR N-PSDB; AAH35367.

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 11; Page 8187-8188; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922

XX Sequence 193 AA;

Query Match 96.4%; Score 27; DB 4; Length 193;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
|||
Db 88 DESGLF 93

RESULT 11

ADP81219
 ID ADP81219 standard; protein; 240 AA.
 XX
 AC ADP81219;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Protein of human ovarian specific gene, SEQ ID NO 253.
 XX
 KW normal; neoplastic; ovarian; ovarian specific nucleic acid; OSNA;
 KW metastatic; cancer; vaccine; cytostatic; human.
 KW
 XX Homo sapiens.
 XX
 PN WO2004053079-A2.
 XX
 PD 24-JUN-2004.
 XX
 XX 08-DEC-2003; 2003WO-US038855.
 XX
 PR 06-DEC-2002; 2002US-0431301P.
 PR 06-DEC-2002; 2002US-0431321P.
 PR 30-JUN-2003; 2003US-0484584P.
 PR 07-NOV-2003; 2003US-0518607P.
 XX
 XX (DIAD-) DIADEXUS INC.
 XX
 PI Macina RA, Turner LR, Sun Y, Liu S, Chen H;
 XX
 DR WPI; 2004-468850/44.
 DR N-PSDB; ADP81060.
 XX
 XX New ovarian specific nucleic acid molecules and polypeptides useful for
 PT diagnosing, preventing or treating ovarian cancer, for producing
 PT transgenic animals or cells, or for research purposes.
 XX
 PS Claim 12; SEQ ID NO 253; 754pp; English.
 XX
 CC The invention relates to novel isolated nucleic acid molecules and
 CC polypeptides present in normal and neoplastic ovarian cells. These
 CC comprise a nucleic acid sequence encoding any of the 167 amino acid
 CC sequences (e.g. 438, 237 or 233 amino acids) fully defined in the
 CC specification (SEQ. ID NOS: ADP81095 to ADP81261) and comprises any of
 CC the 128 nucleotide sequences (e.g. 4798, 1494 or 1691 bp) fully defined
 CC in the specification (SEQ. ID NOS: ADP80967 to ADP81094). The invention
 CC further comprises: a method for determining the presence of a ovarian
 CC specific nucleic acid (OSNA) in a sample; a vector comprising the above
 CC producing a polypeptide encoded by the above nucleic acid molecule; a
 CC polypeptide encoded by the nucleic acid molecule cited above; an antibody
 CC or its fragment that specifically binds to the above polypeptide; a
 CC method for determining the presence of an ovarian specific protein in a
 CC sample; a method for diagnosing or monitoring the presence and metastases
 CC of ovarian cancer in a patient; a kit for detecting a risk of cancer or
 CC presence of cancer in a patient; the kit comprising a means for
 CC determining the presence of the above nucleic acid molecule or
 CC polypeptide; a method of treating a patient with ovarian cancer; and a
 CC vaccine comprising the above polypeptide or nucleic acid encoding the
 CC polypeptide. The isolated nucleic acid molecules and polypeptides have
 CC cytostatic activity. The isolated polypeptides may be used to create a
 CC vaccine. The isolated nucleic acid molecules and polypeptides can be used
 CC for diagnosing or monitoring the presence and metastases of ovarian
 CC cancer and treating ovarian cancer. This sequence represents the protein
 CC of an ovarian specific gene of the invention.
 XX
 SQ Sequence 240 AA;
 Query Match 96.4%; Score 27; DB 8; Length 240;
 Best Local Similarity 83.3%; Pred. No. 2,1e-02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 DESGLF 6
 |||||

Db 100 DESGLF 105

RESULT 12
 ADI43131
 ID ADI43131 standard; protein; 244 AA.
 XX
 AC ADI43131;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Plant transcription factor #541.
 XX
 KW transgenic; plant; enhanced tolerance to abiotic stress;
 KW phosphate tolerance; hormone sensitivity; disease resistance;
 KW sugar sensing; flowering; flower structure; stem bifurcation;
 KW branching pattern; apical dominance; trichome; stem morphology;
 KW root growth; root hair; seed development; cell proliferation;
 KW cell differentiation; premature senescence; necrosis; plant size;
 KW leaf morphology; seed morphology; seed biochemistry; root antracyanin;
 KW plant anthocyanin; light response; shade avoidance; bioinformatic;
 KW transcription factor; gene; ds.
 XX
 OS Oryza sativa.
 XX
 PN US2004019927-A1.
 XX
 PD 29-JAN-2004.
 XX
 PF 25-FEB-2003; 2003US-00374780.
 XX
 PR 18-APR-2001; 2001US-00837944.
 XX
 XX (SHER/) SHERMAN B K.
 PA (RIEC/) RIECHMANN J L.
 PA (JIAN/) JIANG C.
 PA (HEAR/) HEARD J E.
 PA (HAAK/) HAAKE V.
 PA (CREE/) CREELMAN R A.
 PA (RATC/) RATCLIFFE O.
 PA (ADAM/) ADAM L J.
 PA (REUB/) REUBER T L.
 PA (KEDD/) KEDDIE J.
 PA (BROU/) BROUN P E.
 PA (PILG/) PILGRIM M L.
 PA (DUBE/) DUBELL A N.
 PA (YUGG/) YU G.
 XX
 PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
 PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
 XX
 DR WPI; 2004-132245/13.
 XX
 PT New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.
 XX
 PS Claim 1; SEQ ID NO 1594; 435pp; English.
 XX
 CC The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produce a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glyophosphate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure, change in stem bifurcations, altered branching
 CC pattern, reduced apical dominance, reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC altered cell proliferation or cell differentiation; rapid development;

CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins, or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This is the amino acid sequence of a plant
 CC transcription factor, and an orthologue of Arabidopsis thaliana
 CC transcription factors isolated in the invention, that can be used in the
 CC creation of a transgenic plant with altered traits.

XX SQ Sequence 244 AA;

Query Match 96.4%; Score 27; DB 8; Length 244;

Best Local Similarity 83.3%; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DEXGLF 6

Db 26 DEAGLF 31

RESULT 13

AD003227

ID AD003227 standard; protein; 244 AA.

XX AC AD003227;

XX DT 01-JUL-2004 (first entry)

XX DE Thalecress transcription factor, Rice orthologue #126.

XX KW Rice; transcription factor; plant; transgenic; abiotic stress;

XX KW cold tolerance; heat tolerance; drought; osmotic stress;

XX KW phosphate limitation; potassium limitation; nitrogen limitation;

XX KW hormone sensitivity; disease resistance; sugar sensing; seed germination;

XX KW flowering; inflorescence architectural change;

XX KW meristem cell differentiation; phyllotaxy; apical dominance;

XX KW trichome development; seed development; premature senescence;

XX KW delayed senescence; lethality; necrosis; plant size; leaf morphology;

XX KW seed morphology; secondary metabolism; light response; shade avoidance.

OS Oryza sativa.

XX PN US2004045049-A1.

XX PD 04-MAR-2004.

XX PP 10-APR-2003; 2003US-00412699.

XX PR 13-SEP-1999; 99US-00394519.

XX PR 21-JAN-2000; 2000US-00489376.

XX PR 17-FEB-2000; 2000US-00506720.

XX PR 22-MAR-2000; 2000US-00532591.

XX PR 22-MAR-2000; 2000US-00533029.

XX PR 22-MAR-2000; 2000US-00533030.

XX PR 22-MAR-2000; 2000US-00533392.

XX PR 22-MAR-2000; 2000US-00533648.

XX PR 06-APR-2000; 2000WO-US009448.

XX PR 16-NOV-2000; 2000US-00711394.

XX PR 27-MAR-2001; 2001US-00819142.

XX PR 17-APR-2001; 2001US-00837444.

XX PR 30-JAN-2002; 2002US-00959131.

XX PR 14-JUN-2002; 2002US-00171468.

XX PR 09-AUG-2002; 2002US-00225066.

XX PR 09-AUG-2002; 2002US-00225067.

XX PR 09-AUG-2002; 2002US-00225068.

XX PR 17-DEC-2002; 2002US-0434166P.

XX PR 25-FEB-2003; 2003US-00374780.

XX (ZHAN/) ZHANG J.

PA (FROM/) FROMM M E.

PA (HEAR/) HEARD J E.

PA (RIEC/) RIECHMANN J L.

(ADAM/) ADAM L J.
 (BROU/) BROUN P E.
 (PINE/) PINEDA O.
 (REUB/) REUBER T L.
 (KEDD/) KEDDIE J S.
 (YUGG/) YU G.
 (JIANG/) JIANG C.
 (SAMA/) SAWAHA R S.
 (PILG/) PILGRIM M L.
 (CREE/) CREELMAN R A.
 (DUBE/) DUBELL A N.
 (RATC/) RATCLIFFE O.
 (KUMI/) KUMIMOTO R.
 (SHER/) SHERMAN B K.

Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;

Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;

Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;

Sherman BK;

WPI; 2004-225755/21.

New transgenic plant, useful in developing phenotypes with altered or improved characteristics or traits.

Claim 1; SEQ ID NO 1641; 213pp; English.

The invention relates to a transgenic plant comprises a recombinant polynucleotide having a polynucleotide sequence or its complementary sequence comprising a sequence encoding a polypeptide, that initiates transcription (i.e. a transcription factor) from Arabidopsis, Soybean, Rice, Rape or Corn, comprising any of the sequences appearing as AD001588-AD003527 or AD003530-AD003559. Also included are using a transgenic plant to grow a progeny plant, an expression cassette (comprising a constitutive, inducible or tissue-specific promoter and a recombinant polynucleotide described above), a host cell comprising the expression cassette, producing a modified plant having a modified trait, identifying a factor that is modulated by or interacts with a polypeptide encoded by the polynucleotide sequence and identifying at least one downstream the polynucleotide sequence that is subject to a regulatory effect of any of the polypeptides encoded by the polynucleotide described above. The transgenic plant is useful for producing a plant that has an altered trait e.g. an enhanced tolerance to abiotic stress (increased tolerance to chilling, germination in cold conditions, freezing tolerance, tolerance to heat, tolerance to drought, tolerance to osmotic stress, tolerance to salt, tolerance to phosphate limitation, tolerance to potassium limitation, decreased sensitivity to nitrogen limitation), altered hormone sensitivity, reduced sensitivity to abscisic acid, an altered response to ethylene, disease resistance, altered susceptibility to Botrytis, altered susceptibility to Fusarium, altered susceptibility to Erysiphe, altered susceptibility to Pseudomonas syringae, altered susceptibility to Sclerotinia, altered sugar sensing, improved seed germination and seedling vigor, early flowering, late flowering, extended period of flowering, an inflorescence architectural change, a change in stem bifurcations, a lack of a shoot meristem, reduced meristem cell differentiation, altered phyllotaxy, altered branching pattern, reduced apical dominance, reduced trichome density, ecotopic trichome development, altered trichome development, altered stem morphology, increased root growth, increased root hairs, altered seed development, altered cell proliferation/cell differentiation, premature senescence, delayed senescence, lethality, increased necrosis, an increase in seedling or plant size, decreased plant size, a change in leaf morphology, increased altered leaf development, increased leaf size and mass, glossy leaves, leaf cell expansion, change in seed morphology, altered seed coloration, increased seed size, decreased seed size, altered seed shape, change in leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid content, increased leaf insoluble sugars, decreased leaf insoluble sugars, increased leaf anthocyanins, an alteration of leaf fatty acid content, an alteration of leaf glucosinolate content, change in seed biochemistry, an increase in seed oil content, decrease in seed oil content, increase in seed fatty acid content, decrease in seed fatty acid content, increase in seed protein content, decrease in seed protein content, alteration in seed prenyl lipid content, increase in seed

CC sterols, upregulation of genes involved in secondary metabolism, increase
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in
 CC light response or shade avoidance. The present sequence represents an
 CC orthologue of a thalass transcription factor isolated from Rice.
 XX
 SQ Sequence 244 AA;

Query Match 96.4%; Score 27; DB 8; Length 244;
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
 || || ||
 Db 26 DEAGLF 31

RESULT 14
 AAR36397
 ID AAR36397 standard; protein; 265 AA.

XX AC AAR36397;

XX 25-MAR-2003 (revised)
 DT 03-JAN-2003 (revised)
 DT 20-AUG-1993 (first entry)

XX Truncated GA733-2E soluble variant of type I membrane protein.

XX GA733-2F; secretory protein; mutant; cancer; immunotherapy; tumour.

XX Homo sapiens.

XX WO9308298-A1.

XX 29-APR-1993.

XX 08-OCT-1992; 92WO-US008572.

XX 18-OCT-1991; 91US-00778232.

XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX Linnenbach AJ, Koprowski H, Herlyn D;

XX WPI; 1993-152489/18.

XX N-PSDB; AAQ41176.

XX Soluble variants of type I membrane protein, GA733-2E and GA733-2F -
 useful for cancer immunotherapy and detection of tumour cells.

XX Claim 1; Fig 1; 66pp; English.

XX Oligonucleotide primers were synthesised for the PCR synthesis of a
 CC truncated cDNA that encoded the amino terminal signal peptide and
 CC extracellular 212 of the selected antigen, GA733-2. The synthetic primers
 CC correspond to nucleotides 85-114 (sense strand) and 885-919 (antisense
 CC strand) of the native GA733-2 DNA. The primers were used in standard PCR
 CC using GA733-2 cDNA as template. Pharmaceutical compns. contg. GA733-2E
 CC are useful as vaccines for active immunotherapy against carcinomas, and
 CC as reagents for detecting GA733-2 antigens expressed on tumour cells.
 CC (Updated on 09-JAN-2003 to add missing OS field.) (Updated on 25-MAR-2003
 CC to correct PN field.)

XX Sequence 265 AA;

Query Match 96.4%; Score 27; DB 2; Length 265;
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6
 || || ||
 Db 100 DESGLF 105

RESULT 15

ADL23185

ID ADL23185 standard; protein; 265 AA.

XX AC ADL23185;

XX 20-MAY-2004 (first entry)

XX Human soluble Ep-CAM.

XX Human; bactericidal/permeability-increasing protein; BPI; Ep-CAM; CAB2.1;
 KW recombinant polypeptide production; ING-1; antibody; anti-CD18 antibody;
 KW cosmetic product.

XX Homo sapiens.

XX Synthetic.

XX US2003203447-A1.

XX 30-OCT-2003.

XX 31-MAR-2003; 2003US-00404724.

XX 29-MAR-2002; 2002US-0368530P.

XX (HORW/) HORWITZ A H.

XX Horwitz AH;

XX WPI; 2003-875646/81.

XX N-PSDB; ADL23184.

XX Producing recombinant polypeptide, useful for treating or diagnosing
 PT comprises culturing cells transformed or transfected with a vector
 PT comprising multiple copies of a transcription unit separated by a
 PT selective marker gene.

XX Example 12; SEQ ID NO 58; 133pp; English.

XX The invention relates to producing a recombinant polypeptide comprising
 CC culturing cells, which have been transformed or transfected with a
 CC vector, or its segment comprising multiple copies of a transcription unit
 CC separated by at least one selective marker gene, where the transcription
 CC unit encodes a polypeptide under selective conditions. Also included are
 CC a vector or segment comprising multiple copies of a transcription unit
 CC separated by at least one selective marker gene where the transcription
 CC unit encodes a polypeptide, a host cell comprising an expression vector
 CC or segment and a stable cell line comprising an expression vector or
 CC segment. Each transcription unit is under the control of its own promoter
 CC and 3' untranslated region, where the promoter is an SV40, HSV, bovine
 CC growth hormone, thymidine kinase, MPSV, mouse beta globin, human EPI, MSV
 CC -LTR, RSV, MMTV-LTR, CMV, MLV, Chinese hamster elongation factor or mouse
 CC Abelson LTR promoter. The expression vector further comprises multiple
 CC enhancers. The transcription unit also encodes two different subunits of
 CC a multimeric protein, an immunoglobulin light and heavy chain
 CC polypeptides or at least the variable regions of the immunoglobulin light
 CC and heavy chain polypeptides. It further encodes a BPI protein
 CC (bactericidal/permeability-increasing protein) product. The protein
 CC product BPI protein fragment, BPI analogue, BPI variant or BPI-derived
 CC peptide. The transcription unit encodes an RBP21 and is under the
 CC control of an hCMV promoter and mouse light chain 3' untranslated region,
 CC where the vector further comprises 0, 1 or 2 copies of a human heavy
 CC chain enhancer and either a gpt or neo gene. Other genes suitable for
 CC expression using the method of the invention are Ep-CAM and CAB2.1 (both
 CC not defined). The immunoglobulin may be the ING-1 chimaeric mouse/human
 CC antibody (or humanised versions or proline substitution mutants) or an
 CC anti-CD18 antibody. The method is useful for producing recombinant
 CC polypeptide. Recombinant polypeptide compositions are useful in
 CC therapies, in diagnostic procedures or as tools in preventive medicine.
 CC Recombinant polypeptides are also found in a wide array of both health
 CC and cosmetic products, used to increase the quality of life. Complex
 CC polypeptide products are also routinely used in research laboratories

CC both as end products of analyses and as agents in assays for the study or
CC preparation of other molecules. Advantages of the present invention
CC includes increased recombinant polypeptide production, increased
CC production efficiency, greater control and/or regulation over the
CC qualities of the polypeptide expressed, increased stability of cell
CC lines, and/or decreased costs for materials, reagents and/or other
CC resources. The present sequence represents a protein suitable for
CC expression using the method of the invention.
XX

SQ Sequence 265 AA;

Query Match 96.4%; Score 27; DB 7; Length 265;

Best Local Similarity 83.3%; Pred. No. 2.4e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DEXGLF 6

|||

Db 100 DESGLF 105

Search completed: January 3, 2005, 13:49:03

Job time : 158 secs

